

Fusion Proteins Incorporating Lysozyme

The present invention relates to the production of peptides in the milk of transgenic mammals, for example non-human placental mammals.

5

Polymers of amino acids concatenated via their amino and carboxyl groups form the basis for a variety of important biological compounds. Polymers of 3 to 100 amino acids are generally known as peptides, whilst larger polymers are known as proteins. This distinction is purely arbitrary, and polymers of up to about 110 amino acids can still be considered as "peptides". Thus, the term "peptide" as used herein refers to amino acid polymers of 3 to 110 amino acids. Peptides as defined herein may be biologically active without requiring any further modification, or may form the building blocks for larger complex molecules by chemical modification into larger structures or by modification such as glycosylation. The term "peptide" is used herein to include biologically active and inactive polymers, which may or may not have undergone further modification.

15

Peptides have a number of commercial applications, including use as medicaments, nutritional additives and research tools. For this reason, economic, large scale production of peptides is desirable. Direct chemical synthesis of peptides is expensive due to the high cost of reagents and the degree of purification needed to remove failed sequences. Microbial synthesis by recombinant DNA technology is an alternative, but not always appropriate for peptide production due to difficulties in extraction and purification from microbial cells, and the absence of microbial enzymes to perform necessary post-translational modification. Heterologous proteins may be produced in stably transfected mammalian host cell lines, many of which are commercially available today. However, concern remains that these cell lines are derived from tumours of various types.

20

25

As an alternative to the above methods, the production of proteins in the milk of

30

transgenic sheep is possible, as illustrated in WO-A-8800239 and WO-A-9005188. The production of proteins in the milk of transgenic animals has the advantage that large volumes of milk containing the desired protein can be harvested using simple and environmentally safe technology. The use of living organisms to produce proteins means that all the material produced will be identical to the natural product. In terms of amino acid structures, this means that only L isomers will be produced. Also, the number of wrong sequences will be minimised due to the high fidelity of biological synthesis compared to synthetic routes.

- 10 Further, the use of a biological process for the production of the proteins ensures that only biologically safe materials are produced, in contrast to chemical methods where side reactions may produce toxic materials, which can only be removed at additional cost. The use of a biological process also enables some reactions, which are difficult to perform in good yield by chemical means, to be efficiently carried out. For
- 15 example, carboxy terminal amidation of a peptide can be essential for biological activity or for the prolongation of *in vivo* half-life, and is carried out by a specific enzyme which recognises and modifies proteins having a glycine residue at the carboxy terminus (Eipper B. A *et al.*, (1993) *Protein Science* 2 489-497). Therefore, suitably designed proteins produced by means of a transgenic animal will be
- 20 specifically amidated prior to secretion. The amidation of proteins is only one of a number of post-translational modifications which can be carried out by the biosynthetic pathways in the mammary gland and harnessed for the synthesis of biologically active proteins. Other post-translational modifications include disulphide bridge formation, phosphorylation and γ -carboxylation of glutamic acid residues and
- 25 the addition of O- and N- linked glycosylation (Wold, F. *Ann. Rev. Biochem.* 50 783-814).

The technology for the production of large proteins, as opposed to shorter peptides, in large quantities in the milk of transgenic sheep has been well established. For

example, the human protease inhibitor, α_1 -antitrypsin has been produced in the milk of transgenic sheep in excess of thirty grams of protein per litre (Wright, G. *et al.*, (1991) *Bio/Technology*, 9 77-84). It is expected that the same technology can be applied to the production of proteins in cattle, which can produce up to 10,000 litres of milk per lactation.

There are a number of difficulties relating to the secretion of short peptides in mammalian systems due to the nature of the secretory process. Proteins destined for secretion are directed into the endoplasmic reticulum, which forms the first stage of the constitutive secretory pathway, by a short pro-sequence, usually of at least twenty amino acids. The messenger RNA encoding a protein destined for secretion is translated by a ribosome which is initially free in the cytoplasm of the producing cell. However, as the end of the newly synthesised protein emerges from the large ribosome complex, the secretory leader sequence is bound to a 'signal recognition protein' (SRP). The act of binding has two effects. First, it causes the translation and protein synthetic machinery of the ribosome to 'pause' and secondly, it promotes the docking of the ribosome to the surface of the ER. This docking then re-starts translation and the protein destined for secretion is then synthesised through the ER membrane into the inner compartment. During the course of this second synthetic phase, the secretory leader sequence is cleaved off and the protein is folded appropriately. Then, after removal of the secretory leader sequence, and any secondary sequence related to correct folding, by proteolysis, and any other necessary modifications (primary glycosylation events, gamma carboxylation, etc) the protein moves on through the secretory pathway.

The fundamental problem with the secretion of peptides from mammalian systems is the requirement for a secretory leader sequence, the binding of the signal recognition peptide and the geometry of the ribosomal complex. Simple experiments have shown that if ribosomes which are actively translating proteins are treated with a powerful

and non-specific protease, which can degrade all exposed proteins, then sequences of polypeptide about forty amino acids in length are protected. This implies that this sequence is buried within the large ribosomal complex and that only longer sequences capable of binding the SRP will be competent to enter the ER secretory pathway.

5

The requirement for a minimum peptide length was confirmed by studies which truncated normally secreted proteins such as lysozyme (Ibrahim *et al* (1986) *Eur. J. Biochem* **155**(3) 571-6) and insulin (Okun *et al* (1990) *J. Biol Chem.* **265**(13) 7478-84). Shorter versions of lysozyme, which still contained the secretory leader sequence, of 102 and 74 amino acids, were still capable of binding the SRP (as demonstrated by the ability of added SRP to 'pause' translation in a cell-free system) but a 52 amino acid truncation could not. Also, studies on the secretion of truncated insulin confirmed that not only did short peptides not 'reach' the SRP but were also secreted with low efficiency. Therefore, due to the basic mechanism of secretion it is evident that very short peptides cannot enter the secretory pathway. It is also apparent that even if peptides are long enough, with the addition of the secretory leader sequence, to engage the SRP, efficient secretion is unlikely due to a preference for amino acid sequences in excess of perhaps 100 amino acids (Okun *et al* (1990) *J. Biol Chem.* **265**(13) 7478-84). This preference is reflected by the general size of secreted proteins which are normally at least 120 or more amino acids in length. Secretion of peptides shorter than 100 amino acids normally occurs via an entirely different mechanism where peptides are generated by the proteolytic cleavage of larger precursor proteins, sequestered in specialised vesicles within a cell and stored until needed. In this case secretion occurs in response to a specific signal which promotes fusion of the vesicle with the plasma membrane of the cell with concomitant release of the peptide into the external medium.

15

Thus, the basic mechanism by which proteins are secreted, involving ER docking mediated by the SRP, precludes the secretion of very short peptides, of less than perhaps 40 amino acids, and severely decreases the efficiency of peptides less than

30

100 amino acids long. In the absence of a fusion partner, to direct peptides to the secretory pathway, peptides of less than 100 amino acids long are naturally secreted by a completely different vesicle-based mechanism that only operates at high capacity in specialised tissues such as neurones. This pathway does not represent a viable
5 alternative for making peptides in mammary tissue.

A second reason for expressing peptides as fusion proteins in milk is that it is easier to purify a fusion protein from milk, which is a complex biological fluid containing fats, sugars and proteins as well as peptides and proteolytic fragments, than to purify the
10 free peptide. If the properties of the fusion partner dominate those of the peptide, it is likely that at least the initial purification steps will be common to processes for different peptides and thereby reduce development costs for a number of peptides. Regarding purification, the use of a peptide fusion is also beneficial in that two different recovery modalities can be employed: one for the fusion protein and then,
15 after cleavage, one for the peptide. This approach is expected to yield a more pure product, or require fewer stages to achieve higher purity, because peptide impurities will be reduced during the purification of the fusion and protein impurities during the purification of the peptide.

20 The third advantage of expressing a peptide in milk as a fusion rather than as the peptide is that the biological properties of the peptide are likely to be masked and therefore not interfere with the physiology of the host animal. This has been demonstrated for calcitonin where it was shown that the alpha lactalbumin fusion protein was inactive in an in vivo assay designed to measure the depression of plasma
25 calcium levels in the rat in response to an injection. This is in contrast to the cleaved and purified calcitonin which did exhibit biological activity (WO95/27782 and McKee C. *et al*(1998) *Nat. Biotechnol.*

30 The expression of heterologous proteins in mature or fused form in the milk of a transgenic female animal is also described in WO92/22644. This application discloses

fusing a peptide gene sequence into a HINDIII restriction enzyme site in the coding sequence of the WAP gene, in order to express the peptide in milk. This fused gene construct merely serves to target the peptide expression to milk, but does not result in the expression of a fusion protein in milk, and thus is likely to suffer from the above mentioned problems of the art.

WO 95/27782 describes processes for the production of peptides in the milk of transgenic animals based on expressing the peptide linked to a "fusion partner protein". The fusion protein can be isolated from the milk and subsequently cleaved to release the desired peptide. In a preferred embodiment the use of human α -lactalbumin as a fusion partner protein linked to calcitonin as the desired peptide, is described. Human α -lactalbumin is a small, natural milk protein capable of terminal extension, thus satisfying some criteria of a fusion partner protein. However, it has demonstrated that human α -lactalbumin fusion constructs are expressed in the milk of rabbits at only 2.1mg/ml (PCT/GB95/00769; McKee C. *et al* (1998) *Nat. Biotechnol.* **16**(7) 647-651), a low yield compared to the yield of non-fusion α_1 -antitrypsin at 30 grams per litre. It is to this problem of low expression of such fusion proteins that the present invention is addressed.

Thus, in a first aspect of the present invention there is provided a process for the production of a peptide, the process comprising expressing in the milk of a transgenic non-human placental mammal a fusion protein comprising the peptide to be expressed linked to a fusion partner protein which is lysozyme. Suitably, the process also includes the steps of separating the fusion protein from the milk, and cleaving the fusion protein to yield the peptide.

Lysozyme is a natural milk protein. It has been found to satisfy all the essential criteria of an ideal fusion partner protein whilst, surprisingly, enabling high yields of peptide compared to those achieved with other natural, milk-derived fusion partner

- proteins. Lysozyme is a small molecule of approximately 14,000 Daltons mass and containing about 120 amino acids, depending on the species. Therefore, the mass yield of peptide linked to lysozyme as the fusion partner protein will be high per mole of fusion protein produced. Due to its untypically high content of basic amino acids,
- 5 lysozyme is simple and inexpensive to purify from expression media, such as milk. A further desirable feature of lysozyme, and an essential feature of any fusion partner protein, is its ability to carry an amino- or carboxy- terminal extension without any substantial effect on expression level or structural stability.
- 10 Lysozyme (EC 3.2.1.7) is a naturally occurring protein, secreted into bodily fluids, such as milk, saliva and airway secretions of a number of eukaryotic and prokaryotic species. Lysozymes are 1,4- β -N-acetylmuramidases which act as anti-bacterial agents by hydrolyzing the glycoside bond between the C-1 of N-acetylmuramic acid and the C-4 of N-acetylglucosamine in bacterial peptidoglycan. As mentioned above,
- 15 these enzymes are small (M_r 14,000 – 15,000) and basic in nature (pH 9.5-11). Lysozymes have been widely studied, as is apparent from hen egg white lysozyme which was the first protein containing all of the 20 common amino acids to be sequenced, the first enzyme for which a 3-D model was deduced by X-ray crystallography and the first enzyme for which a detailed mechanism of action was
- 20 provided (Fischer *et al.*, (1993), *Applied. Microbiology. and Biotechnology.* 39:537-540).
- The lysozyme fusion partner protein of the present invention may be from any mammal which naturally secretes lysozyme in bodily fluids. The preferred lysozymes
- 25 are those which are expressed at more than 5g/l in the milk of transgenic animals and are stable with carboxy terminal extensions. Preferably, the lysozyme is that of a placental mammal, for example humans, cattle, sheep, goats, rabbits and rats. The lysozyme of other animals, such as chickens, may also be useful. It is also desirable

that the genomic DNA functions well behind the beta-lactoglobulin promoter and is stable when built into DNA constructs.

The present invention is based upon the surprising observation that much higher levels of expression can be achieved by using lysozyme as a fusion partner protein than, for example, α -lactalbumin, a small protein related to lysozyme with respect to sequence and disulphide bridges. The effectiveness of lysozyme in the expression system of the invention is seen, not only in terms of the absolute expression level of fusion protein, which is forty fold higher than that observed with α -lactalbumin, but also in the increase in the proportion of animals expressing fusion proteins at high levels. This latter observation has the surprising advantage that less G_0 founder animals are needed to generate a high expressing transgenic line. Although there is a possibility that generating and screening hundreds of α -lactalbumin founders would eventually give high expression levels, this is not viable commercially, especially since large animals such as cattle, sheep and goats which are preferred for the present invention, are expensive to generate. Indeed, in view of the much higher expression levels achieved with the lysozyme fusion protein, and the inclusion of an 'insulator' to moderate integration position effects, it seems unlikely such expression levels could ever be achieved with the α -lactalbumin fusion protein.

The improvement in expression of the fusion protein with lysozyme over α -lactalbumin is observed in experiments where both DNA fusion constructs use the β -lactoglobulin promoter driving genomic constructs, thus excluding the possibility that the higher expression levels are attributable to differences in regulatory sequences. Likewise, in similar experiments, drastically improved expression was seen with lysozyme over α -lactalbumin, even after the inclusion in each construct of an 'insulator' sequence which is designed to isolate an integrated transgene, or transgene array, from effects of 'position' within the host genome (US Patent No. 5, 610, 053). Further experiments have shown that the lysozyme fusion protein expression is

superior even when there is no difference in target peptide or cleavage site between the fusion proteins.

- As mentioned above, lysozyme as a fusion partner protein is expressed at a higher level but, as importantly, the expression is consistently high for a number of transgenic lines. For instance, as seen in Example 1 the level of expression of the lysozyme-cyanogen bromide-calcitonin fusion protein in mice is 4.3 ± 3.5 (n=19) with the highest level being 11.0mg/ml, in contrast to the α -lactalbumin-cyanogen bromide- β -lactoglobulin construct where the average expression level is much lower, at only 0.16 ± 0.10 (n=12), with the highest level being 0.26 mg/ml. Although there is not necessarily a direct correlation between the number of copies of a transgene integrated into the host genome and the resulting expression level, because of the very strong effect of integration site on expression, the disparity in expression between the two fusion partners cannot be attributed to poor integration of the α -lactalbumin construct.
- Copy number measurements for the above examples for expression levels in mice show that the copy numbers are between 5 and 30 copies for at least half of the lines analysed with either of the fusion partners.

- Part of the rationale for using a fusion partner is that the expression level of the target peptide should be less dependent on the properties of the individual peptide. It is therefore expected that high level expression and high frequency of good expression seen with lysozyme should extend to any peptide. This is confirmed by experiments in mice using a lysozyme fusion with peptide GLP-1. Here, the same high expression lines and frequency of high expressing levels were seen in agreement with the results for the lysozyme-calcitonin fusion. The average expression level in mice was 15.4 ± 5.4 (n = 4) with the highest level being over 22.9 mg/ml. These results surprisingly show an even greater improvement over the results with calcitonin even though less animal lines were analysed.

- The unexpected high level expression levels associated with the lysozyme fusion protein is not limited to particular transgenic species. Rather, the high expression levels seen in mice, compared to α -lactalbumin fusion proteins, is also seen in rabbits. As previously described, the enterokinase cleavable α -lactalbumin – calcitonin fusion
- 5 protein was expressed in rabbit milk at 1.55 ± 0.45 mg/ml ($n=3$) with a maximum level of 2.1mg/ml. In contrast, the lysozyme-GLP-1 fusion, again with the enterokinase cleavable linker, was expressed at 12.2 ± 9.3 mg/ml ($n=9$) in rabbit milk, with a maximum expression level of 31.0 mg/ml.
- 10 The difference in expression efficiencies between lysozyme and α -lactalbumin is unexpected, in light of previous work on transgenic expression. A published report of bovine -lactalbumin expression from a cDNA in mouse milk gives an expression level of 0.45mg/ml. (Vilotte, J.L. *et al*, (1989) *Eur J Biochem* Dec 8; 186 (1-2): 43-48). This allows a direct comparison with the expression of a human lysozyme cDNA
- 15 construct expressed in mouse milk when the maximum expression was 0.71mg/ml (Maga, E.A. *et al*, (1995)) and supports the premise that there is no intrinsic property in either molecule which predisposes a comparatively high expression level of lysozyme over α -lactalbumin.
- 20 In the fusion protein expression studies, both the lysozyme and α -lactalbumin genomic sequence constructs are driven by the same β -lactoglobulin promoter and carry the same 3' untranslated region, also from the β -lactoglobulin gene. A second possible explanation for the high lysozyme expression levels is that there is some factor, for example an enhancer of transcription, which is present within the genomic lysozyme
- 25 sequence, which is absent from that of α -lactalbumin. If this were so, it would be predicted that the same comparatively higher lysozyme levels would be seen in natural human milk when compared to the levels of α -lactalbumin. This is not the case and, in fact, higher levels of α -lactalbumin (about 3 grams per litre of milk) are observed

(Lonnerdal B *et al.* (1976) *Am. J. Clin. Nutr.* **29**(10) 1127-1133) compared to lysozyme (0.25 grams per litre) (Goldman A.S *et al.* (1982) *J. Pediatr.* **100**(4) 563-7).

Thus, there is no evidence, from either expression studies of cDNA of lysozyme or -lactalbumin in transgenic milk or from the behaviour of the regulatory sequences driving expression in milk, to suggest that the expression level for lysozyme would be substantially higher than that of β -lactalbumin when both are expressed off genomic constructs using the same promoter and 3' environments.

Peptides produced by the present invention are preferably from 3 to 110, preferably 3 to 100 amino acids in length, but the invention is not limited to the production of peptides of the preferred range. The invention is particularly suitable for producing peptides which require post-translational modification in order to be biologically active, or improve *in vivo* half life, for example α -amidation. Many peptides found in the nervous and endocrine system of animals and bioactive peptides from other sources which have actions on the nervous system are -amidated. Examples include:

α -amidated residue

	A alanine	b,o CRH; p Galanin; μ -Conotoxin
20	C cysteine	crustacean cardioactive peptide; conotoxins G1, M1, SI
	D aspartic	deltorphin
	E glutamic	joining peptide
	F phenylalanine	FMRF-NH ₂ ; gastrin; cholecystokinin; CGRP; γ MSH
25	G glycine	oxytocin; vasopressin; GnRH; pancreastatin; leucokinin I, II; <i>Manduca</i> adipokinetic hormone; leucokinin I, II
	H histidine	Apamin; scorpion toxin II
	I isoleucine	h,r CRH; PHI; <i>Manduca</i> diuretic hormone; rat neuropeptide EI (melanin concentrating hormone)
	K lysine	ELH; cecropin A; PACAP38, conotoxin GIA
30	L leucine	b,h GHRH; b-amidorphin; mastoparan; cecropin B; buccalin; myomodulin; PACAP27; proglucagon (111-123)

	M methionine	Substance P; Substance K; PHM; gastrin releasing peptide; neurokinin A,B; neuromedin B, C
	N asparagine	VIP (mammalian); neuromedin U; corazonin; mast cell degranulating peptide
	P proline	calcitonin; TRH
5	Q glutamine	melittin; levetide
	R arginine	preproglucagon (89-118)
	S serine	frog granulin-R
	T threonine	rat galanin; avian VIP; locust adipokinetic hormone
	V valine	α MSH; r.p.h secretin; metorphamide/adrenorphin
10	W tryptophan	cockroach myoactive peptide, sea anemone peptide; crustacean erythrophore concentrating peptide
	Y tyrosine	NPY; PYY; PP; ω -conotoxin; amylin

15 *PACAP, pituitary adenylate cyclase activating peptide.

An example of a biologically active peptide which is of medical and commercial interest is calcitonin. Other examples of peptides include parathyroid hormone, glucagon, glucagon-like-peptide-1, and members of the general classes of peptide: 20 magainins, histatins, protegrins and clavainins. Calcitonin, for example, is a 32 amino acid peptide which contains a single disulphide bridge and is amidated at the carboxy terminus. The peptide hormone is secreted by the thyroid or parathyroid gland in mammals and by the ultimobranchial bodies in other vertebrates and serves to lower the level of calcium in the blood by reducing the level of release of calcium from 25 bone. It is a highly functionally conserved molecule, and the protein obtained from salmon has widespread therapeutic applications for example in the treatment of Paget's disease, hypercalcaemic shock and osteoporosis.

Another peptide of potential commercial interest is glucagon-like-peptide-1 (GLP-1). 30 This is a 30 amino acid carboxy-terminal amidated peptide which is secreted by both gut cells and within the hypothalamus in response to feeding. Its main action is to potentiate the glucose stimulation of insulin secretion and to help regulate gastric

emptying. It is therefore being evaluated as a potential therapy in the treatment of diabetes.

5 In addition to the above examples of specific peptides, there is an entire class of peptides that have anti-microbial activity, which will be required in large quantities and are therefore especially suitable candidates for production using the transgenic fusion protein approach. These peptides work by disrupting biologically important membranes usually by the creation of ion-permeant pores. Many of these are amidated and this modification possibly functions to increase biological half-life, by preventing
10 degradation by carboxy peptidases, and may also be important in reducing the net negative charge of the peptide, by modifying the acidic carboxylic group. Examples of antimicrobial and cytotoxic peptides include those belonging to the classes magainins, histatins, protegrins and clavainins.

15 A fusion protein produced by the above noted methods forms a second aspect of the invention.

When lysozyme is used as a fusion partner protein, it may be appropriate to add to the carboxy-terminus an extension which serves as a linker to join the fusion partner
20 protein to the peptide. The linker is at least 10, 15 or preferably at least 20 amino acids in length. This is the first demonstration that a large (greater than 20 amino acids) carboxy-terminal extension can be expressed on lysozyme at high levels and without disrupting the stability of the fusion partner. Although the linker may consist of any sequence of amino acids, in order to reduce any adverse effects of the linker on
25 the structural stability of the fusion protein, it is preferred that the linker has neutral structural properties, for example a neutral pH and small sized amino acids. A preferred carboxy terminal extension is flexible linker having the sequence (gly-gly-gly-gly-ser)₃ (SEQ ID NO 1). The provision of a fusion protein comprising a fusion partner protein and peptide joined by means of a flexible linker having the sequence

(gly-gly-gly-gly-ser)₃ represents a third aspect of the present invention. Preferably, the fusion partner protein of the third aspect is lysozyme.

Apart from the presence of any carboxy-terminal linker sequence on the lysozyme fusion partner protein, there may be some variation in the sequence of the lysozyme from a natural sequence. Although natural, wild-type sequences of lysozyme are preferred, some variation from the natural sequence may be accommodated or, in some cases at least, desired, provided that the properties of lysozyme are not compromised to an unacceptable degree. Amino acid homology of at least 90% or 95% will be appropriate and generally not more than 2 or 3 amino acid changes will be preferred. Homology is determined by standard programs such as BLAST provided by the National Centre for Biotechnology Information (<http://www.ncbi.nlm.nih.gov>).

Lysozyme is basic in character. It has an isoelectric point of 10 to 11, and thus carries a positive charge up to this pH range. These characteristics have been exploited to purify lysozyme from other proteins, using techniques such as ion exchange chromatography and affinity chromatography. For example, cation exchange chromatography and affinity chromatography have been used in the purification of human saliva lysozyme (Vasstrad *et al.*, (1980) *Scandinavian Journal of Dental Research* **88** 219-228). Cation exchange chromatography has also been used in the purification of human airway lysozyme (Jacquet *et al.*, (1987) *Analytical Biochemistry* **160**: 227-232) and in the rapid purification of lysozyme from hen egg white (McCreath *et al.*, (1997) *J. Chromatography A* **773**: 73-83). The purification of lysozyme can also be accomplished using anion exchange chromatography in a negative mode. This technique has been used by in the purification of hen egg white lysozyme (Vachier and Awade (1995) *J. Chromatography B* **664**: 201-210). In this case, hen egg white was diluted with a buffer at pH 9.0 and applied to a column of Q Sepharose FF (Amersham Pharmacia Biotechnology). At this pH most of the egg proteins had a net negative charge while the lysozyme still possessed a net positive charge and was either not retained or weakly bound to the column, thereby facilitating

its purification. More sophisticated techniques for the purification of Lysozyme from a number of sources have also been described. For example, hydrophobic interaction chromatography following a cation exchange capture step has been used in the purification of lysozyme from horse milk, by exploiting a Ca^{2+} dependant change in the hydrophobic/hydrophilic nature of Lysozyme which mediates its interaction with the hydrophobic resin (Noppe *et al.*, *Journal of Chromatography A* 719: 327-331).

In the purification of bacteriophage lysozymes, two approaches have been taken. For the purification of phage lambda lysozyme expressed in *E.coli*, a negative purification step on an anion exchange resin (DEAE-cellulose) followed by a positive purification on a cation exchange resin (S-Sepharose Fast Flow, Amersham Pharmacia Biotechnology) was used (Jespers *et al.*, (1991) *Protein Engineering* 4: 485-492). An alternative method whereby lysozyme is expressed with a poly-Histidine tag, thereby allowing its purification using immobilised metal ion affinity chromatography has also been used (During, (1993), *Protein Expression and Purification* 4: 412-416; Sloane *et al.*, (1996) *J. of Biotechnology* 49: 231-238). This latter method for the purification of lysozyme requires the use of an affinity partner to aid its purification. It is surprising to note that the present invention is based upon the use of lysozyme as an ideal fusion partner to aid purification, whereas previous work discussed above suggests that lysozyme is a protein which itself requires an affinity partner for purification.

According to a preferred feature of the first aspect of the present invention, any suitable method may be used for purification of the fusion protein from the expression media. Preferably, a method for the purification of lysozyme will be used for initial and final purification. For example, precipitation techniques could be used. Lollike, K. *et al.*, (Leukemia 9 : 206-209, 1995) have described the purification of human lysozyme from neutrophils using a combination of PEG precipitation and column chromatography. The solubility characteristics of lysozyme have been studied in some depth (Curtis, R.A. *et al.*, *Biotechnol. Bioeng.* 57: 11-21, 1998) and it has been recognised that solubility characteristics can be influenced by anion binding.

Precipitation with ammonium sulphate has also been used in the part purification of human milk proteins including Lysozyme (Brignon, G & Ribadeau-Dumas, B., *Biochimie* 64: 231-235, 1982), again combined with chromatography; in this case, a size exclusion column. Size exclusion chromatography (SEC) may be considered a relatively high resolving technique for Lysozyme-peptide fusions due to their relatively small size. Following on from this, SEC could also be used as a way of separating the Lysozyme molecule from the peptide after cleavage. A new precipitation technique using carbon dioxide has also proved useful in the fractional precipitation of Lysozyme from protein mixtures (Winters, M.A. *et al.*, *Biotechnol. Bioeng.* 62: 247-258. 1999), and may be applicable to the present invention.

Another technique which discriminates on the basis of size and that could be used for Lysozyme purification is filtration, in particular tangential flow filtration as disclosed in PCT WO 97/42835. Using this technique it may be possible to separate Lysozyme-peptides fusions or lysozyme from milk or other process fluids. It is further apparent that techniques such as either ion exchange chromatography or affinity exchange chromatography may be used. These methods are generally applicable and inexpensive, and can be used in the purification of lysozyme-peptide fusions from milk. In one such procedure, milk at a suitable pH and ionic strength could be applied to either a packed, fluidised or expanded bed for direct capture of the lysozyme fusion peptide. The majority of milk proteins being acidic in nature would be expected to pass through the column or to bind weakly to the resin, whereas the basic lysozyme fusion protein would bind tightly to the column. High purity material should result from elution with an increase in either ionic strength, solution conductivity, pH or a combination of any or all three. The use of a fluidised bed chromatography column for the purification of lysozyme from milk has been described recently in the literature (Noppe *et al.* *Journal of Chromatography A* 719: 327-331). Alternatively another procedure may be adopted by which a "negative" purification is used. In this application, milk is added to the anion exchanger using any of the column methods above whereby depending on the pH, acidic protein will become bound, whereas the

lysozyme fusion protein would not bind or only bind weakly. If required, further purification of the lysozyme-peptide fusion protein may be carried out using either hydrophobic interaction chromatography or affinity chromatography, as discussed above.

5

It is therefore apparent that Lysozyme is a protein that lends itself to purification using a number of differing process operations thus increasing its attractiveness as a fusion partner. A general scheme for the purification of a lysozyme-peptide fusion may therefore be presented as a combination of low cost precipitation and/or filtration techniques optionally followed by the use of a number of column chromatographic techniques including, but not exclusively limited to, cation exchange chromatography, anion exchange chromatography, size exclusion chromatography and affinity chromatography. Although it would be of significant economic benefit if only low cost purification techniques, such as precipitation and filtration, are used, it may be expected that in most cases, the use of high resolution chromatographic techniques may be used to provide the high purity required for most pharmaceutical applications. Additionally, these chromatographic techniques would be useful in the elimination of possible pathogen such as viral particles. The applicability of any general purification scheme, once established, can than be validated for a number of different lysozyme fusions and, where necessary, changes in the unit operations or operating conditions can be made.

Once purified, the peptide can be released from the fusion protein by any suitable means. In order to achieve cleavage of the peptide, there is preferably provided a cleavage site between the fusion partner protein and peptide to enable release of the peptide from the fusion protein. Preferably, this is in addition to the flexible linker. Any suitable cleavage site may be used, including for example those which are cleaved by chemical or enzymatic means. An example of the former is treatment with cyanogen bromide which breaks peptide bonds at the carboxyl side of a methionine residue. The advantage of this method is that the reaction uses inexpensive reagents,

30

but an important restriction is that there can be no internal methionine in the target peptide otherwise this too will be cut. However, methionine is a low abundance amino acid and many peptides of potential interest as commercial targets will satisfy this criterion. A second possible complication with cyanogen bromide may arise if the fusion partner also contains a methionine residue, since this could result in cleavage within the fusion partner, releasing a peptide which can interfere with the purification of the target peptide and adversely impact on production costs. This may be circumvented by using species of lysozyme which does not carry an internal methionine, for example human. Therefore, a fusion protein made from human lysozyme and salmon calcitonin, which also lacks an internal methionine, and preferably also a flexible linker would be a suitable candidate for cleavage by cyanogen bromide. In the case of salmon calcitonin, which has an amino cysteine residue, efficient cyanogen bromide cleavage requires the prior sulphonation of the adjacent thiol. This prevents an irreversible side reaction, and the thiol can be regenerated after the cleavage reaction is complete (Ray M. V. L *et al* (1993) *Bio/Technology* **11** 64-70). A variety of other chemical cleavage reactions are also possible and any of these could be applied to the appropriately designed fusion protein (Han K. K *et al.*, (1983) *Int. J. Biochem* **15** 875-884).

A second preferred method of cleaving the fusion protein to release the peptide is to link the carboxy terminus of the fusion partner protein to the peptide via a sequence of amino acids which includes a specific recognition site for enzymatic cleavage, and which does not occur anywhere else in the fusion protein. Examples of such sites are the sequences Ile-Glu-Gly-Arg (SEQ ID NO. 2) and Asp-Asp-Asp-Lys (SEQ ID NO. 3), which are recognised and cleaved by blood factor Xa and enteokinase respectively. This approach has the advantage that the cleavage enzyme can be chosen by reference to its recognition sequence; certain enzyme recognition sequences only occur very rarely in natural molecules, such as those quoted above. However, if at least part of the cleavage sequence occurs naturally at appropriate ends of the peptide or the fusion partner protein, then that fact can be used in the practice of the invention.

Where a linker is also present between the fusion partner protein and the peptide, in accordance with the second aspect of the invention, then the linker may be reduced or omitted as appropriate.

5 The cleavage site may contain more sequence than is absolutely necessary to direct cleavage. For example, in the case of a cleavage site recognised by enterokinase, the activation peptide of trypsinogen, the natural substitute cleaved by enterokinase, may be included as part of the linker. This has the sequence Phe-Pro-Thr-Asp-Asp-Asp-Lys.

10

It is important to note that the combination of the linker region, such as (gly-gly-gly-gly-ser)₃, with either the cyanogen bromide cleavage site or the enterokinase cleavage recognition sequence still results in a high level expression of peptide. Part of this success in achieving high levels of peptide expression when using a fusion protein including a flexible linker in combination with a range of different cleavable sequences can be attributed to the neutral structural properties of the linker.

15

After cleavage, the peptide can be easily separated from the fusion partner protein by any convenient method. In the case of lysozyme, the efficient removal of the redundant fusion partner protein can be achieved by exploitation of the basic character of the protein, using for example ion exchange or affinity chromatography, as discussed above.

20

In the practice of the present invention, fusion proteins are produced in the milk of transgenic animals. The design and production of DNA sequences which encode the protein-peptide fusion proteins is well known to those skilled in the art (Sambrook *et al.*, Molecular Cloning- A Laboratory manual, Cold Spring Harbor Laboratory Press, (2nd Edition) 1989). The lysozyme coding sequence can be obtained by screening libraries of genomic material or reverse translated messenger RNA derived from the animal of choice. These sequences are then cloned into an appropriate plasmid vector

25

30

and amplified in a suitable host organism, usually *E. coli*. The DNA sequence encoding the peptide would then be constructed, for example by polymerase chain reaction amplification of a mixture of overlapping annealed oligonucleotides. If the production of a carboxy-terminal amino peptide was the objective then a glycine codon would also be introduced at the 3' terminus of the sequence encoding the peptide. This material would then be joined to the 3' end of the DNA encoding the lysozyme with the inclusion of a linker sequence, preferably according to the second aspect of the invention, including an appropriate fusion protein cleavage site. This entire construct after checking that the desired sequence has been constructed would be cloned into a suitable vector carrying control sequences suitable for the generation of transgenic animals.

After amplification of the vector, the DNA construct would be excised with the appropriate 3' and 5' control sequences, purified away from the remains of the vector and used to produce transgenic animals. Conversely, with some vectors, such as yeast artificial chromosomes (YACS), it is not necessary to remove the assembled vector; in such cases the vector may be used directly to make transgenic animals.

According to a fourth aspect of the present invention, there is provided an isolated or recombinant DNA molecule encoding a fusion protein, the DNA sequence comprising a coding sequence having a first segment encoding a fusion partner protein which is lysozyme coupled to a second segment encoding a peptide. Suitably, the coding sequence may be operatively linked to a control sequence, which enables the coding sequence to be expressed in the milk of a transgenic non-human placental animal.

To enable release of the peptide it is further desirable for the DNA molecule to include a sequence, provided between the first and second segments, which encodes a cleavage site, as discussed above in relation to the first aspect. Further, a flexible linker sequence may also be provided, with or without the cleavage site between the

first and second segments. In a preferred feature, the linker has the sequence (gly-gly-gly-ser)₃.

- 5 A DNA sequence which is suitable for directing the production to the milk of transgenic animals carries a 5' promoter region derived from a naturally derived milk protein and is consequently under the control of hormonal and tissue specific factors. Such a promoter is therefore most active in lactating mammary tissue. This promoter may be followed by a (usually shorter) DNA sequence directing the production of a protein leader sequence which would direct the secretion of the fusion protein across the epithelium into the milk. At the other end of the fusion protein construct a suitable 3' sequence, preferably also derived from a naturally occurring milk protein may be added. The 3' sequence performs various poorly defined functions and is to improve the stability of the transcribed RNA and thus increase the levels of the protein. An example of suitable control sequences for the production of proteins in the milk of transgenic animals are those derived from ovine β -lactoglobulin; see for example, WO-A-8800239 and WO-A-9005188, which describe these control sequences in particular, and more generally address the production of transgenic animals secreting proteins of interest in their milk.
- 10
- 15
- 20 The DNA molecules of the invention can conveniently form part of vectors, suitable for use in transforming host cells, either for the direct production of the fusion protein, or alternatively, for use in the production of a transgenic organism, preferably a non-human placental mammal, wherein the fusion protein is obtainable from the milk of said mammal. Suitable vectors include plasmids, such as those described herein, but other forms of vectors can be used and the skilled person will be aware of these and how they may be used/manipulated. Thus, such vectors form a sixth aspect of the invention. Host cells transformed with such vectors form a further aspect of the invention. Suitably, the host cells are mammalian.
- 25

Therefore, according to a yet further aspect of the present invention, there is provided a transgenic non-human placental mammal whose genome incorporates a DNA molecule comprising a coding sequence having a first segment encoding a fusion partner protein which is lysozyme coupled to a second segment encoding a peptide.

5 The DNA molecule may also be modified as described above to incorporate such features as the control sequences, linker sequence and/or a cleavage site.

The production of transgenic animals can now be performed using a variety of methods. The most common of these is pronuclear injection where the DNA, having

10 first been purified away from vector sequences, is directly injected into the male pronucleus. This can be done with either genomic sequences of cDNA constructs co-injected with genomic sequence for an endogenous milk protein (Clark, A. J *et al* (1992) *Bio/Technology* 10 1450-1454; and WO-A-9211358). Examples of other methods include cytoplasmic injection into ova, transformation of totipotent stem cells

15 or carriage of foreign DNA sequences by sperm (Pursel, V. G and Rexroad, Jr C. E (1993) *J. Anim. Sci* 71 (Suppl. 3) 10-19). A wide variety of animals are suitable for transgenic expression in milk, including cows, sheep, goats, rabbits, mice and pigs. Essentially, any species which is domesticated and produces sufficient quantities of harvestable milk would be preferable for the production of lysozyme-peptide fusion

20 proteins.

In a final aspect of the present invention, there is provided a composition comprising a fusion protein of the invention. The composition may be expression media derived from a transgenic, non-human placental animal, and preferably is milk.

25 Preferred features of each aspect of the present invention are as for each other aspect, *mutatis mutandis*.

The invention will now be illustrated by the following Examples, with reference to the

30 figures, in which:

FIGURE 1 shows the structure of pCLYSM construct.

FIGURE 2 shows the DNA sequence of pCLYSM, excluding the bacterial plasmid (SEQ ID NO.4).

5

FIGURE 3 shows the results of mass analysis of pCLYSM fusion purified from mouse milk.

Criteria Used In Hypermass Calculation:

10 Agent , Mass: 1.0079, Charge: 1, Agent Gained

Charge Estimation Tolerance: 0.1000

Tolerance Between Mass Estimates: 20.0000

	Peak	Intensity	Predicted Peak	Charge	Hypermass Estimate
15	1404.13	94062.50	1404.13	14.00000	19643.73
	1512.03	61250.00	1511.62	13.00353	19643.35
	1637.84	33750.00	1636.35	12.01093	19641.98
	1786.54	28750.00	1785.37	11.00723	19640.90
20	Final Estimated Mass: 19642.49				
	Standard Deviation: 1.30				

FIGURE 4 shows the monitoring of cyanogen Bromide cleavage and sCT re-folding by mass analysis.

25

FIGURE 5 is a diagrammatic illustration of the lysozyme-linker-enetrokinase-GLP-1 fusion construct.

30	Species	Predicted Mass
	Fusion-Gly	19968 Da

Fusion amidated	19911 Da
GLP-1 7-37	3355 Da
GLP-1 7-36 amide	3298 Da

- 5 FIGURE 6 shows SDS-PAGE analysis of GLP-1 fusion milks.

Lanes 1 and 8 Partially purified GLP-1 Fusion

Lanes 2 and 9 Control rabbit milk

10	Lane 3	GLP-1 Line 2	Lane 10	GLP-1 Line 50
	Lane 4	GLP-1 Line 20	Lane 11	GLP-1 Line 61
	Lane 5	GLP-1 Line 21	Lane 12	GLP-1 Line 64
	Lane 6	GLP-1 Line 38	Lane 13	GLP-1 Line 66
	Lane 7	GLP-1 Line 46		

15

FIGURE 7 shows a western blot analysis of selected GLP-1 fusion milks.

	Lanes 1 and 6	Control Rabbit Milk
	Lanes 2 and 7	GLP Line 2
20	Lanes 3 and 8	GLP Line 21
	Lanes 4 and 9	GLP Line 50
	Lanes 5 and 10	GLP Line 64

FIGURE 8 shows ESI-MS analysis of GLP-1 fusion milks.

25

FIGURE 9 shows an 8DS-PAGE gel (Coomassie blue stained) of milk obtained from the sheep expressing the lysozyme – enterokinase cleavable linker – calcitonin fusion protein, as discussed in Example 3.

FIGURE 10 shows a Western blot of the fusion protein of Example 3 after a single cation exchange purification step.

5 Example 1 - A comparison of expression levels in mice using α -lactalbumin and lysozyme fusion proteins both joined to calcitonin via a cyanogen bromide cleavable linker.

10 Two constructs were designed to express calcitonin fusion proteins. The first termed *pCALM*, was designed to express a human alpha-lactalbumin / salmon calcitonin fusion protein in the milk of transgenic animals. This fusion protein allows the release of calcitonin from the end of a linker arm fused to the alpha-lactalbumin C terminal by cyanogen bromide (CNBr) chemical cleavage.

15 *pCALM structure*

15 *pCALM* consists of :

1. A 4.2kb region comprising the ovine b-lactoglobulin (BLG) promoter and 5' untranslated region (UTR).
2. A 2069bp region comprising the complete coding region of the human alpha lactalbumin gene corresponding to bases 750 to 2819 of the human alpha lactalbumin sequence (EMBL + Genbank database, accession number : X05153), previously
20 derived from a library of cloned human genomic DNA in bacteriophage lambda.
3. A 162bp region encoding a C terminal extension to the alpha-lactalbumin protein, comprising a (Gly₄ Ser)₃ Ala Ser linker arm, CnBr cleavage site and salmon calcitonin peptide sequence extended by a single Gly residue to facilitate C-terminal
25 amidation and translational stop signal.
4. A 2.5kb region comprising the 3' UTR, polyadenylation site and 3' flanking region of the ovine beta-lactoglobulin gene.
5. A 242 bp region comprising the chick b-globin insulator region (US patent 5,610,053).

6. The pUC18 bacterial plasmid vector.

Further details about some of the components of *pCALM* are described in McKee C, *et al.*, *Nat Biotechnol* 1998 Jul;16(7):647-51

A second construct, termed pCLYSM, was designed to express a human lysozyme - salmon calcitonin fusion protein in the milk of transgenic animals. This fusion protein allows the release of calcitonin from the end of a linker arm fused to the lysozyme C terminal by cyanogen bromide chemical cleavage.

pCLYSM structure

The structure of pCLYSM is shown in Figure 1, and the sequence in Figure 2. pCLYSM consists of:

1. A 4.2 kb region comprising the ovine β -lactoglobulin (BLG) promoter and 5' untranslated region (UTR).
2. A 4.8 kb region comprising the complete coding region of the human lysozyme gene corresponding to bases 520 to 5345 of the human lysozyme sequence (EMBL+ Genbank database, accession number: X14008), derived by polymerase chain reaction (PCR) amplification of genomic DNA prepared from the human cell line HT1080.
3. A 162 bp region encoding a C terminal extension to the lysozyme protein, comprising a (Gly₄ Ser)₃ Ala Ser linker arm, CNBr cleavage site and salmon calcitonin peptide sequence extended by a single Gly residue to facilitate C-terminal amidation.
4. A 2.5 kb region comprising the 3' UTR, polyadenylation site and 3' flanking region of the ovine BLG gene.
5. A 242bp region comprising the chick α -globin insulator region (US patent 5,610,053)
6. The pUC18 bacterial plasmid vector.

Although the insulator region was included in this construct it was not essential to the production of the fusion protein.

5 *Expression of pCALM and pCLYSM in transgenic mice*

pCALM and pCLYSM were introduced into mouse zygotes by pronuclear microinjection using standard procedures (Palmiter *et al.*, (1982) *Nature* 300: 611-615). Mice containing the pCALM or the pCLYSM transgenes were identified by
10 PCR analysis and Southern blotting. Milk was obtained from these mice and analyzed by a combination of SDS-PAGE, western blotting and RID.

Copy Numbers and Expression Data for pCALM mouse lines.

Line	Copy No.	Expression Level mg/ml
1.40	20	0.26
1.7	30	0.13
2.22	20	0.13
3.18	2	0.26
3.27	2	0.13
3.38	1	0.026
3.4	1	0.026
5.1	7	0.26
5.7	10	0.26
6.1	7	0.26
7.1	1	0.013
8.18	3	0.13

Copy Numbers and Expression Levels of pCLYSM Mouse Lines

Line	Copy No.	Expression Level mg/ml
156.6	10	9.8
156.6.1	>30	5.4
156.6.2	>30	11
156.6.3	10-30	9.8
159.11	1	0.55
159.3	10	0.13
161.17	10	2.75
161.2	1	1
162.10	5	4.74
162.13	1	1.20
162.15	1	0.4
162.8	5	6.4
162.8.1	1	5.4
162.8.2	1	7.5
162.8.3	1	7.5
163.13	30	2.23
163.3	20	5
164.18	2	1.09

5

Characterisation of the lysozyme fusion protein.

In view of the high expression levels of the lysozyme fusion protein it is possible to purify sufficient material from mouse milk for further characterisation and cyanogen

bromide cleavage. The pCLYSM protein product was purified from EDTA solubilized milk by cation exchange chromatography and further characterized by mass analysis. The observed mass of 19642.5Da is consistent with the full length fusion protein containing five disulphide bridges, four in lysozyme and one in sCT (Figure 3).

5

Prior to CNBr cleavage the cysteine residues were sulphonated according to the method of Ray *et al* (Biotechnology Vol.11 Jan 1993). Briefly, the fusion protein was incubated at pH 8.0 in the dark for 12hrs with a 10 fold molar excess of sodium sulphite and a 2 fold molar excess of sodium tetrathionate. Following desalting, the
10 sulphonated protein was cleaved in 8M urea, 50mM HCl pH8.0 containing 1mg CNBr/mg protein. Cleavage was monitored by mass analysis and final refolding of sCT achieved by 10 fold dilution of the reaction mixture into 0.1M Tris/Cl pH8.0, 10mM cysteine (Figure 4)

15 Example 2.- Production of a human lysozyme – glucagon-like peptide 1 (GLP-1) in the milk of transgenic rabbits.

Details of the GLP-1 construct:

A construct termed pGLUC-1 was designed to express a human lysozyme / human
20 glucagon-like peptide 1 GLP1 fusion protein in the milk of transgenic animals. This fusion protein allows the release of GLP1 from the end of a linker arm fused to the lysozyme C terminal by enterokinase enzymic cleavage.

pGLUC1 structure

25

The structure of pGLUC1 is identical to pCLYSM, as described in example 1, apart from the C terminal extension to the lysozyme protein. pGLUC1 consists of:

1. A 4.2 kb region comprising the ovine b-lactoglobulin (BLG) promoter and 5' untranslated region (UTR).

2. A 4.8 kb region comprising the complete coding region of the human lysozyme gene corresponding to bases 519 to 5344 of the human lysozyme sequence (EMBL+ Genbank database, accession number: X14008), derived by polymerase chain reaction (PCR) amplification of genomic DNA prepared from the human cell line HT1080.
3. A 162 bp region encoding a C terminal extension to the lysozyme protein, comprising a [(Gly4 Ser)₃ Ala Ser] linker arm, a (Asp Asp Asp Asp Lys) enterokinase cleavage site, a 30 amino acid region corresponding to residues 7-36 of the human GLP1 sequence followed by a single Gly residue to facilitate C-terminal amidation and a translational stop signal.
4. A 2.5 kb region comprising the 3' UTR, polyadenylation site and 3' flanking region of the ovine BLG gene.
5. A 242bp region comprising the chick b-globin insulator region (US patent 5,610,053)
6. The pUC18 bacterial plasmid vector.

15

A construct coding for human lysozyme, a flexible enterokinase cleavable linker and GLP 7-37 was microinjected into the nucleus of recently fertilised rabbit embryos. The last glycine residue of the peptide was expected to act as a suitable substrate for the amidating enzyme, PAM. The embryos were transplanted into recipient does and the resultant kits screened for the transgene by PCR and Southern analysis. A total of 12 transgenic offspring were produced (3 females and 9 males). All offspring were induced to lactate with oxytocin injections, the females at two months and the males at three. Milk samples were collected for five days or until the rabbits dried off, which ever was the shorter.

25

2. Milk analysis

2.1 Human Lysozyme Radial Immuno Diffusion analysis

All milks were analysed on Human Lysozyme RID plates obtained from the Binding site. Control rabbit milk produced no signal at 1 in 100 dilution (in assay diluent supplied with the kit) whereas the transgenic milks produced precipitin rings from 1 in 250 to 1 in 4000. The expression levels by RID are as follows:

Rabbit No.	Sex	Milk Volume (mls)	Expression (mg/ml)
2	M	10.75	17.7
20	M	8	0.9
21	M	20.7	16.6
38	M	31.75	3.75
46	M	13.5	4.47
50	M	4.35	17.5
60	F	Drops	>40
61	F	13.9	14.6
64	M	1.3	31.0
66	F	0.25	3.05

RID values were measured against human lysozyme, assuming that GLP-1 fusion and lysozyme have the same response. To validate this the Coomassie Blue staining intensity for GLP-1 milks and lysozyme standards were compared – the RID and Coomassie Blue estimates were in broad agreement.

2.2 SDS-PAGE

Milk from each founder and a non-transgenic control rabbit was diluted in reducing Laemmli buffer and run at the equivalent of 0.1 μ l on a 4-20% Novex gel. The gel was visualised with Coomassie Blue stain (Figure 6). Although rabbit milk has a band that co-migrates with the GLP-1 fusion, it is clear that the high expressing milks have a

much more intense band in this position. Rabbit 60 is not included because it is deemed to be a failed induction and not representative of a natural lactation.

2.3 Amino-Terminal Sequence Analysis

5

In order to confirm the identity of the over-expressed protein N-terminal sequence analysis was carried out. 1µl of milk from Male No.2 was run on a 4-20% Novex gel as described above and electroblotted to PVDF membrane. The membrane was stained briefly with Coomassie Blue and the fusion band cut out for Edman sequencing. A strong sequence corresponding to that of human lysozyme was observed. The co-migrating band was not identified, although it is presumably a rabbit casein which is not in the Swiss Prot database.

10

2.4 Western Blotting for Human Lysozyme and GLP-1

15

In order to further confirm the identity of the fusion protein Western blotting was done with both anti lysozyme and anti GLP-1 antibodies. Four high expressing milks (2, 21, 50 and 64) and control rabbit milk were run in duplicate at 0.01µl on 4-20% Novex gels and transferred to PVDF membrane for western blotting. Membranes were blocked with 2% BSA in PBS, 0.1 % Tween 20 and then probed with Dako Rabbit anti human lysozyme or GLP-1 Mab. The blots were visualised with ECL reagent from Pierce and Amersham Hyperfilm. While control rabbit milk gave no signal, a single band which cross-reacts with both antibodies was observed in each transgenic milk (Figure 7). Milk 64 alone gave rise to two minor bands, one above and one below the main signal – the identity of these species is not yet known.

20

25

2.5 GLP Fusion Purification Purification

Lysozyme was chosen as a fusion partner because of its basicity and resultant ease of purification from the main acidic proteins in milk by cation exchange chromatography.

GLP-1 milks were mixed with an equal volume of 200mM EDTA pH8.0 (to dissociate the casein micelles) and then diluted 10 fold in 20mM Tris/Cl pH7.0. This feedstock was bound on a Pharmacia Mono S column equilibrated in the same buffer and eluted with a gradient of 0 - 0.5M NaCl. A single peak containing GLP-1 fusion and associated caseins eluted early in the gradient. Although not pure this material was suitable for analysis by Electro Spray Ionisation-Mass Spectroscopy (ESI-MS).

2.6 ESI-MS Analysis and Enterokinase Cleavage

ESI-MS analysis was performed using a API100 mass spectrometer (Perkin Elmer). Samples were applied using Micro Protein Cartridges, 1.0x10mm C8, washed with 0.1% acetic acid and eluted with 0.1% acetic acid in acetonitrile. Mass calibration was performed using Polypropyleneglycol standards.

Mass analysis revealed two species, the major one being consistent with amidated GLP-1 fusion with conversion of the C-terminal Gly to Arg-amide, the second minor species 80Da higher (Figure 8). Treatment of the fusion with alkaline phosphatase confirmed that the 80Da increase was due to phosphorylation, a modification found in other milk proteins. Incubation of the fusion with enterokinase followed by mass analysis revealed that the phosphate group was on the fusion partner not GLP-1 and confirmed the release of fully amidated peptide (Figure 8). The analysis described above was applied to lines 2, 20, 21, 38, 61 and 66. All showed the same pattern of phosphorylation and were also fully amidated.

3. Summary

A total of 10 transgenic founders for the GLP-1 fusion were produced. All lines were found to express the fusion protein as judged by human lysozyme RID, SDS-PAGE and western blotting analysis at between 0.9 and 31mg/ml. The fusion was easily purified by cation exchange chromatography, allowing characterisation by ESI-MS

and cleavage with enterokinase to yield GLP-1 7-36amide. Processing of the glycine extension by endogenous PAM activity was very efficient and no residual glycine extended fusion was detected.

Example 3 - Analysis of the expression in sheep of a lysozyme fusion protein joined to calcitonin via an enterokinase cleavable linker.

A construct, designated CALC11, was designed for the expression of a calcitonin fusion protein in the milk of transgenic sheep. This fusion protein allows the release of calcitonin from the end of the linker arm fused to lysozyme by incubation with the protease enterokinase.

The structure of the CALC11 construct is identical to that described for pCLYSM except that the CNBr cleavage site (methionine) is replaced by an enterokinase cleavage site (aspartic acid – aspartic acid – aspartic acid – aspartic acid – lysine).

A CALC11 transgenic sheep was generated by nuclear transfer as described in WO 97/07669 and WO 97/07668. Milk obtained from induced lactation of this animal was analysed by a combination of SDS-PAGE and Western blotting.

Figure 9 shows an SDS-PAGE gel (Coomassie Blue stained) of milk obtained from the sheep expressing the lysozyme-enterokinase cleavable linker-calcitonin fusion protein.

Lanes 1 and 6 - molecular weight markers

Lanes 2 and 3 - Control sheep milk

Lanes 4 and 5 - lysozyme-calcitonin fusion milk

The fusion protein is clearly visible in transgenic milk (arrows), compared to control milk. The second new band, below the fusion protein, is a clipped fusion protein

where part of the carboxy terminus has been removed during the secretory process by an unidentified protease.

The level of expression of the lysozyme-enterokinase cleavable linker-calcitonin fusion protein in the sheep milk, based on a gel comparison of the purified fusion protein and a known quantity of lysozyme, is 5 mg/ml.

Figure 10 shows a Western blot of the fusion protein after a single cation exchange purification step.

Lanes 1, 2 - purified lysozyme-enterokinase cleavable linker-calcitonin fusion protein probed with anti-calcitonin antibody

Lane 3 - purified α -lactalbumin-calcitonin fusion protein probed with anti-calcitonin antibody (used as a control to show specificity of the calcitonin antibody). The production and characterisation of this fusion protein is described elsewhere (McKee C. *et al* (1998) *Nat. Biotechnol.* 16(7) 647-651).

Lanes 4 and 5 purified lysozyme-enterokinase cleavable linker-calcitonin fusion protein probed with anti-lysozyme antibody.

Probing with the antibody to calcitonin confirms the presence of the calcitonin peptide as part of the fusion protein and shows that, as expected, the size of the lysozyme fusion is similar to that of the α lactalbumin calcitonin fusion used as a standard. The presence of the lysozyme component of the fusion protein is also confirmed by blotting with the antibody to lysozyme. Note the absence of the lower band of the doublet as probed with the antibody to calcitonin confirming that the second band on the SDS-Page gel shown above is indeed truncated before the peptide portion. However, this clipped version is a minor component and does not detract from the observation that full size lysozyme calcitonin fusion has been expressed at a high level in the milk of a transgenic sheep.